Medical Libraries and Computers

The Role of Medical Libraries in Medical Informatics

NINA W. MATHESON, ML, *Baltimore*

The classic function of health sciences libraries is to build and maintain a knowledge base and to provide timely access to that collective memory for the purposes of learning, teaching, caring for patients, conducting research or managing an organization. The formats and representation of that knowledge base are changing rapidly, as are the methods and techniques for gaining access to information. Medical libraries have long used computers for cataloging and controlling records but are now shifting to acquiring, managing and distributing bibliographic and full-text information to local library "networks."

(Matheson NW: Medical libraries and computers—The role of medical libraries in medical informatics, *In* Medical informatics [Special Issue]. West J Med 1986 Dec; 145:859-863)

A world encyclopedia no longer presents itself to a modern imagination as a row of volumes printed and published once and for all, but as a sort of mental clearinghouse for the mind, a depot where knowledge and ideas are received, sorted, summarized, digested, clarified and compared. It would be in continual correspondence with every university, every research institution, every competent discussion, every survey, every statistical bureau in the world. . . . This Encyclopedic organization need not be concentrated now in one place: it might have the form of a network . . . it would constitute the material beginning of a real world brain.[1]

. . . It would be well not to underestimate the potential of technology. . . . whatever one man can dream, sooner or later another man can build.[2]

The potential of medical informatics is limited only by our imaginations. In time, a rational system for obtaining new information, relating it to existing knowledge, organizing it to support decision making and making it available quickly and usefully anywhere in the world will exist—and this information system, like Wells's world brain, will be our libraries. Such libraries will come to exist because of scientific imperatives, not because of technologic developments. Science does not advance by piling up information but by organizing and compressing it[3]—formidable intellectual tasks in aid of which technology has few tools to offer. These libraries may come to exist more quickly in the biomedical sciences than others, however, because of the technologic intensiveness of these disciplines and because of the leadership at the federal level by the National Library of Medicine (NLM), the National Institutes of Health and the National Science Foundation, among others. A significant barrier to rapid progress lies in the limited number of scientists trained in the disciplines of medical informatics.

The need to switch information, not documents, and to actively process information was forcefully expressed in two major documents more than 20 years ago. In 1963 the President's Science Advisory Committee stated, "The ultimate aim is to connect the user, quickly and efficiently, to the proper information and only the proper information."[4] Reporting to the Council of Library Resources, Licklider called for a meld of library and computer into a "procognitive system" that "will not only present information to people but also process it for them, following procedures they specify, apply, monitor and, if necessary, revise and reapply."[5]

The comprehensive electronic libraries of the future are likely to take longer than the next decade to develop because some very difficult intellectual problems must be solved in a number of crucial areas: *knowledge representation*—the expression of knowledge in symbolic form; *knowledge and data acquisition*—methods and techniques for building and updating knowledge bases; *cognitive processing*—the integration of knowledge and strategies of problem solving, and the *human and machine interface*—methods that allow humans and machines to be mutually comprehensible.

Using available technologies over the past century, libraries have created functional systems to manage these problems at the physical materials level. Over the past two decades libraries have concentrated on converting these manual record systems. National classification systems, cataloging codes, subject-heading dictionaries, indexing and abstracting

From the William H. Welch Medical Library and the Department of Medical Information, Johns Hopkins University School of Medicine, Baltimore.
Reprint requests to Nina W. Matheson, MD, Johns Hopkins University, 1900 E Monument St, Baltimore, MD 21205.

ABBREVIATIONS USED IN TEXT

IAIMS = Integrated Academic Information Management System
ILS = Integrated Library System
NLM = National Library of Medicine
OCLC = On-line Center for Library Cataloging
UMLS = Unified Medical Language System

systems, interlibrary loan networks and union catalogs now exist in electronic form. Health sciences libraries have been in the forefront of these efforts, but the breakthroughs needed to manage actual information and knowledge depend on new, interdisciplinary research efforts in medical informatics.

In the meantime, libraries continue as eclectic reference centers and sources of knowledge in a number of formats—paper books and journals, audiotapes and videotapes, computer software, electronic textbooks, journals and data bases and so forth—that are organized to provide needed information in a timely fashion. That libraries and information services as presently constituted are underutilized is a problem librarians have struggled for decades to overcome. Use is almost exclusively determined by accessibility, and accessibility is a function of distance and effort. Studies show that library use is an inverse function of distance.[6] The same is true for communication. Regular and consistent communication decreases dramatically when people are separated by more than 30 m (100 ft). The "principle of least effort"[7] governs the use of services, not the quality of such services. In fact, Allen has predicted that "improving the quality or performance of a particular information service will not lead to increased use of the service."[6(p 186)] Consequently, the focus on mechanisms that provide easy and timely access must be foremost.

The speed, relevance and delivery mechanisms should improve dramatically over the next decade with advanced technology. These improvements will occur for at least three reasons: (1) publishers are moving towards distributing books and journals in electronic formats at the same time that personal computer use is intensifying, (2) libraries will purchase these new formats and train people to use them effectively, (3) librarians will collaborate with medical informatics researchers to develop first-stage electronic libraries that are comprehensive, integrated, information service systems supporting intellectual work.

The concept of a comprehensive, integrated information service system is emerging under different names: in businesses as information resource management,[8] in universities as Project Athena and Intermedia,[9,10] in research libraries as a comprehensive electronic service systems[11] and in health sciences libraries as Integrated Academic Information Management Systems (IAIMSs).[12] The facilitating technologies for using these systems are sophisticated and powerful workstations, integrated through networks with many different systems that retrieve and manipulate information.[13]

Computers in Academic Health Sciences Libraries

One excellent perspective on the development of computer systems in health sciences libraries is available,[14] but a full-scale review has yet to be written. In this brief overview some of the uses of computer technologies by academic health sciences libraries over the past 25 years are highlighted. No attempt is made to chronicle the extraordinary leadership and creativity of the National Library of Medicine or to trace the emergence of electronic bibliographic data bases or to discuss automation in general and university research libraries in particular, for which there are excellent sources.[15-17] My purpose is to describe the foundations of the next decade's developments.

There are three basic uses of computers in today's libraries. The first is to manage and control their collections of books, journals and other information resources. The second is for data-base searching and development. The third is for networking and providing information.

Library Management and Control Systems

Library management and control systems are analogous to business systems and hospital information systems. Providing ready access to a library's resources, whether large or small, requires efficient record control. On a daily basis libraries carry out thousands of individual transactions in buying and receiving materials, maintaining an inventory record, tracking the status of materials at all times and accounting for services rendered. Table 1 provides some indication of the size and scale of a year's activity in academic health sciences libraries.

Academic health sciences libraries began to use computers for library record control systems in the early 1960s. Two pioneering libraries—the Washington University School of Medicine Library under the direction of Estelle Brodman and the UCLA Biomedical Library under the direction of Louise Darling—began with serial control systems. Over the next 20 years, what started as single-function systems to record the receipt of individual journal issues evolved by 1981 into full-scale multifunction, integrated library control systems—the Bibliographic Access and Control System at Washington University (St Louis)[19] and ORION at UCLA.[20] "Integrated library systems" (ILSs) permit a number of functions to be done within a single system based on a master unit record. Thus, the record of a book purchase becomes the basis for the accounting, the catalog and the circulation record.

In the 1970s, thanks to standardization of bibliographic data records under the Library of Congress's machine-readable cataloging format and the arrival of on-line computer technology, libraries were able to share bibliographic data nationally, thereby eliminating the need for every library to create its own records de novo. The On-line Center for Library Cataloging (OCLC) system was operating in 1971 and

TABLE 1.—*Summary Operational Statistics of US Medical School Libraries 1984-1985 (N=126)**

Categories	High	Low	Median	Aggregate Total
Clients	15,000	344	3,976	531,520
Print volumes in collection	619,422	11,468	140,403	19,841,187
Print volumes added annually	16,241	643	5,573	744,537
Serial titles	6,765	390	2,100	295,215
In/out traffic	789,380	10,590	238,156	28,911,103
Collection uses	844,457	7,213	232,218	26,198,306
Interlibrary loans	42,887	257	8,501	311,932

*From the Annual Statistics of Medical School Libraries in the United States and Canada, 8th Ed.[18]

available through national dial access by 1974. By February 1976, 2 million records were in the system; 10 years later 13 million records are available.[21] Now when a library purchases a book, 90% of the time the book will already have been catalogued by another library and the record is available for reuse with little or no editing. Initially, libraries used the OCLC system to produce printed catalog cards and stored the digital records for later use. Thus, when local minicomputer-based library control systems began to appear in the early 1980s, many libraries were well positioned to replace the traditional card catalog with an on-line system.

The National Library of Medicine's ILS software provided the springboards for the full-scale automation of health sciences libraries that is happening today.[22] Like the NLM's earlier Teletypewriter Exchange Network of the Abridged *Index Medicus* experiment, which was instrumental in launching the on-line data-base industry, the ILS public-domain software spawned a new industry that has made it possible for medium-sized libraries to automate quickly and relatively inexpensively. The leading systems among health sciences libraries are the OCLC LS2000[23] and the Georgetown University Medical Center (Washington, DC) Library Information System.[24] In 1983 only a handful of the 126 medical school libraries were significantly automated and able to provide on-line access to their catalog records. By the end of 1987, nearly 40% will have achieved a significant level of automation. Due to the large size of university research libraries, their library control systems have developed differently and more slowly than health sciences libraries.

Data-Base Searching Services

When electronic data bases in medicine, science, engineering and technology emerged in the 1960s, there were no more than a few dozen bibliographic files. By 1976 there were more than 300 publicly available for on-line searching. In 1984 more than 2,800 electronic data bases offered electronic journals, textbooks, newspapers, numeric data bases and bibliographic files in virtually all disciplines through the services of more than 360 vendors worldwide. A handful of vendors provide the primary data bases in biomedicine at hourly prices ranging from $16 to $300 a connect hour.[25] The number of records contained in data bases has grown exponentially from 52 million to 1.68 billion in the past decade. (*ASIS* [American Society for Information Sciences] *Bulletin* 1986; 12:2).

Data-base searching has gone through phases mirroring the development of technology and its diffusion. Searching has moved from a centralized batch mode to a centralized on-line mode and on to a decentralized on-line mode, and has shifted from an institutional service to a personal activity. Before 1971 data-base searching was a batched process conducted at only a few sites. Libraries formulated search requests for their patrons, mailing them to a search center that did the coding and processing and mailed the results to the requester. The whole process usually took two weeks. In 1971 health sciences librarians formulated search queries and carried them out on line through direct access to the NLM computer. Between 1976 and 1984 librarians did most of the on-line searching using a handful of vendors. Users required search intermediaries for a number of reasons: access was not simple; search protocols are complex and subject to continual change; on-line charges are very expensive, and few people had experience with computer terminals, personal computers or telecommunications systems.

Today, now that access is easier and more people are familiar with the procedure, libraries are finding that more and more users prefer to do their own bibliographic searches.[26,27] Many use microcomputer software to do automatic log-on and transfer the search results into local memory units. Libraries are responding to this trend in a number of ways: first, by offering training in effective search techniques, and, second, by beginning to secure electronic data bases for on-premises use, either for local on-line network access or through compact-disk technologies.

Health sciences professionals exhibit a consistent information-seeking pattern, especially for keeping up-to-date. They skim a few authoritative general journals and the primary journals of their subspecialties. Capitalizing on this behavior, vendors such as Bibliographic Retrieval Service/Saunders Colleague and MEDIS are offering a number of journals for full-text on-line retrieval. Some of the major titles are *The New England Journal of Medicine, The Lancet, The British Medical Journal, Nature, The Journal of the American Medical Association* and nine other American Medical Association journals.[28,29] Whether this medium offers an adequate solution to the problem of keeping up-to-date with quality information is a question that has yet to be answered.[30–33]

Today's data bases will not be adequate for long. Knowledge will only continue to expand exponentially. By delivering bulk information more rapidly, they accentuate the need for information processing through selection and compression.

Library Networking

Thanks to the leadership of the National Library of Medicine, virtually all health sciences libraries, no matter how small, are part of a regional medical library network. The principal purpose of the network has been to facilitate access to the literature through an interlibrary loan system of participating resource libraries. Each year about 2 million loans of photocopies of articles or original works flow swiftly through the system.[34]

The communications technologies used have evolved from paper mailed forms, through the use of the Teletypewriter Exchange Network to current electronic mail systems. Printed union lists of journal title holdings have given way to a sophisticated electronic request routing system called DOCLINE. DOCLINE automatically sends a request from a library to the nearest library in the system holding the journal title. If the designated library is unable to fulfill the request for any reason, the request is automatically forwarded through a chain of libraries until it is either successfully completed or reported as unfillable.

DOCLINE is extraordinarily useful and will continue as the backbone of the national interlibrary loan system until superseded by readily accessible electronic journals. The system depends on libraries reporting their holdings in a timely fashion. Two additional features need to be developed. Libraries must be able to report holdings to the journal issue level if the system is to work optimally. Also, a way must be found to couple the very swift and efficient electronic location of a document with effective and inexpensive electronic de-

livery of the same, not only to the originating library but to the person who requested the document in the first place. Efficient and cost-effective facsimile-transmitting technology is urgently needed.

The library-to-library telecommunications linkages are being built rapidly. The electronic linkage between the library and its institutional users is yet to be created in most medical centers. It is towards this goal that the National Library of Medicine's IAIMS program is directed. The concept is to develop interconnections between and among many disparate files, such as patient medical records, the biomedical literature, clinical and research laboratory data, hospital information systems, drug information systems and so forth so that at any stage of decision making information can be drawn from relevant files. Significant economic, technologic and behavioral barriers confront those who hope to achieve this first-level linkage. Local telecommunications networks must be laid, the network architecture must accommodate heterogeneous hardware and software systems, file security must be assured, information files must be designed for effective use in an on-line problem-solving environment, the user community must accept and adhere to a set of system standards, training must be carried out and end-user technology must be ubiquitous and part of the fabric of everyday activity. Essentially, a network culture must emerge.

What Have Medical Libraries to Do With Medical Informatics?

For many in the field of biomedical library science, what role medical libraries play in the area of medical informatics is the burning question of the day and for the future. The answers to the question are not at all clear. The traditional domain of libraries has been the medical literature. Bibliographic indexes and catalogs have been the primary focus. Some libraries are moving into the business of electronically storing and retrieving the medical literature itself, as they have the paper formats. Further roles are on the horizon, however. Schoolman has suggested that if libraries "are to be relevant in supporting clinical decision making, they must be prepared to deal with a variety of new information attributes." Further, he says "they will have to develop access to and the ability to synthesize patient records and hospital management information systems."[35] Even more provocative is the idea that libraries should be responsible for the acquisition, storage and retrieval of the working papers exchanged by scientists exclusively over electronic network communications, which in some fast-moving fields is the equivalent to publication.[36]

Whatever the content of files, it is the storage, retrieval and delivery issues in medical informatics that are likely to involve librarians. For the near term, efforts may be concentrated on developing more intelligent means of access to bibliographic and factual data bases. Libraries will provide gateways to searching multiple external information data bases and they will acquire and manage bibliographic data bases, such as MEDLINE, the Institute of Scientific Information's *Current Contents* and specialized factual data bases, such as those on drug information, for distribution through local area networks. In some instances libraries will independently or collaboratively develop on-line data bases to meet local needs, such as creating directories or special subject bibliographies. Other libraries will begin to develop software that facilitates the use of many data bases and that links library-operated systems to users through networks or direct dial access.

These functions mean that some libraries will need considerably more computer power and new staff with technical computing skills to add to the expertise of librarians. Additional costs are to be expected in information processing. In a review of the field of hospital information management, for example, Grams and co-workers forecast "continued growth in the data processing expenditures to the area of 6% to even 10% of the total annual budget of an institution."[37] This commitment of new resources to libraries is a necessary one. Vendors are intent on securing a market for their own files under proprietary searching software and not on optimizing the use of a competitor's data bases. Current search software, for example, is woefully difficult to use due to a lack of self-evident procedures in moving users between data bases easily and in providing feedback mechanisms to improve products. It will be up to entrepreneurial libraries and librarians to build tools that meet the information needs of health care professionals.

These are the first, admittedly primitive and necessary although not sufficient, steps toward the ultimate and distant goal of knowledge represented in a way that is susceptible to machine inference—that is, new knowledge acquisition, assimilation and manipulation similar to Wells's concept of a world brain.

A world brain requires a world language or at the least the ability to understand and translate all languages. A critical and fundamental task for applying research on medical informatics is to develop the logical models and structure necessary to create a unified medical language. The Unified Medical Language System (UMLS) envisioned by the NLM in its bold and visionary program is "a single classification structure to which all medical language can be mapped" ("Request for Proposal NLM-86-111/PSP," *Commerce Business Daily,* Apr 11, 1986, PSA-9066, p 2). The tasks to be accomplished under the UMLS program are multidimensional and complex. They include characterizing the language of different information sources (literature, clinical records, data banks and knowledge bases) and designing methods of linking, merging and integrating existing thesauruses and classification schemes. The array of skills and knowledge required to work in this domain includes medical sciences, medical informatics, information science, linguistics, artificial intelligence, decision making, medical terminology and thesaurus construction.

Libraries and librarians can contribute to this major undertaking in a number of ways. Librarians who specialize in information services are experts in retrieval languages. This expertise frequently requires an understanding of the system structure and a thorough knowledge of query languages, thesauruses and classification systems. In addition, a skilled reference librarian, like a physician, is a diagnostician—able to elicit the precise nature of the information needed and to select the proper source from dozens that will satisfy this need. Expert systems that can thread their way through a labyrinth of information sources to the answer need to be built by librarians. Finally, iterative testing and evaluation of these experimental applications requires an appropriate "laboratory."

The library that manages a system of data bases that is integrated into an institutional information system is an ideal laboratory.

When the distinctions between communication and publication are less obvious and the boundaries between medical data and medical knowledge blur, the functions of health sciences libraries may be to manage and deliver synthesized knowledge.

Summary

The focus of computer usage in medical libraries is shifting from controlling, processing and managing library records to managing and distributing bibliographic and full-text data-base information on premises through local networks. These telecommunications networks will facilitate the exchange of information from disparate sources in a continuous electronic mode. It will be possible for users at their desks to locate information from an array of geographically distributed data bases, to transfer it to their local and personal workstations, to reformat and manipulate it, to assimilate it into another intellectual form and to redistribute it as a new document, publication or data base. Working with those in the field of medical informatics, libraries will assist in the developing, testing and dispersing of new mechanisms for storing and processing knowledge.

REFERENCES

1. Wells HG: World Brain. London, Methuen, 1938

2. Shaw RR: Implications for library services. Libr Q 1955; 25:344-355

3. Simon HA: Administrative Behavior, 3d ed. New York, Free Press, 1976

4. President's Science Advisory Committee (Weinberg AM, Chair): Report: Science, Government, and Information: The Responsibilities of the Technical Community and the Government in the Transfer of Information. Government Printing Office, 1963

5. Licklider JCR: Libraries of the Future. Cambridge, Mass, MIT Press, 1965

6. Allen TJ: Managing the Flow of Technology. Cambridge, Mass, MIT Press, 1977

7. Zipf GK: Human Behavior and the Principle of Least Effort. Cambridge, Mass, Addison-Wesley, 1949

8. Henderson JA: An IRM manager's manifesto. Inf Resources Management 1986; 1:9-22

9. The 'Star Wars' universities: Carnegie-Mellon, Brown, and MIT. Curr Issues Higher Educ 1983-84; 2:3-24

10. Yankelovich N, Meyrowitz N, van Dam A: Reading and writing the electronic book. Computer 1985 Oct; 18:15-29

11. Grosch AN: The challenge of the eighties—the comprehensive electronic service system (CESS). Electronic Libr 1983 Apr; 1:117-140

12. Planning for Integrated Academic Information Management Systems: Proceedings of a symposium sponsored by the National Library of Medicine, Oct 17, 1984, US Dept of Health and Human Services (DHHS), Public Health Services. Bethesda, Md, Government Printing Office, 1985

13. Crechine JP: The next generation of personal computers. EDUCOM Bull 1986; 21:10

14. Pizer IH: Looking backward, 1984-1959: 25 years of library automation—A personal view. Bull Med Libr Assoc 1984 Oct; 72:335-348

15. Miles WD: A History of the National Library of Medicine, the Nation's Treasury of Medical Knowledge, USDHHS, Public Health Service. Government Printing Office, 1982

16. Adams S: Medical Bibliography in an Age of Discontinuity. Chicago, Medical Library Assoc, 1981

17. Grosch AN: Distributed Computing and the Electronic Library: Micros to Superminis. White Plains, NY, Knowledge Industry Publications, 1985

18. Annual Statistics of Medical School Libraries in the United States and Canada, 8th Ed. Houston, Assoc of Academic Health Sciences Library Directors and the Houston Academy of Medicine, Univ of Texas Medical Center Library, 1986

19. Kelly EA, Yedlin DK, Crawford SY, et al: On-line integrated library system: Bibliographic Access and Control System of Washington University School of Medicine. Bull Med Libr Assoc 1982; 70:281-288

20. Werner G: Managing an academic library, pt 1. J Libr Admin 1985; 6:33-43

21. The OCLC System: A chronology of innovation. OCLC Newslett 1986; 161:13-15

22. Goldstein CM: Integrated library systems. Bull Med Libr Assoc 1983; 71:308-311

23. Lutz M: OCLC's local library system: LS/2000. Libr Hi Tech 1984; 2:7-16

24. Broering NC: The Georgetown University Library Information System (LIS): A minicomputer-based integrated library system. Bull Med Libr Assoc 1983; 71:317-323

25. Williams ME: Electronic databases. Science 1985; 228:445-456

26. Horowitz GL, Jackson JD, Bleich HL: PaperChase: Self-service bibliographic retrieval. JAMA 1983; 250:2494-2499

27. Sohn N, Robbins RD: Computer-assisted surgery (Letter). N Engl J Med 1985; 312:924

28. Clancy S: BRS/Saunders Colleague: An information service for medical professionals. Database 1985 Jun, pp 108-121

29. Collen MF, Flagle CD: Full-text medical literature retrieval by computer. JAMA 1985 Nov 15; 254:2768-2774

30. Blair DC, Maron ME: An evaluation of retrieval effectiveness for a full-text document retrieval system. Communications ACM 1985; 28:289-299

31. Amadio PC: The retrieval of medical literature (Letter). JAMA 1986; 255:2160

32. Channin DS: The retrieval of medical literature (Letter). JAMA 1986; 255:2159

33. Dabney DP: The retrieval of medical literature (Letter). JAMA 1986; 255:2158

34. U.S. National Library of Medicine: Programs and Services FY 1984. Bethesda, Md, NLM, 1985

35. Schoolman HM: The Role of the Medical Library in Support of Clinical Decision Making. Presented at a symposium on Medical Education in the Information Age, sponsored by the Association of American Medical Colleges, Washington, DC, Mar 1985

36. Turner JA: Scholars weigh library's role in collecting computerized research data. Chronicle Higher Educ 1986; 32:34

37. Grams RR, Peck GC, Massey JK, et al: Review of hospital data processing in the US (1982-1984). J Med Syst 1985; 9:175-269